DOCUMENT RESUME

ED 040 406 CG 005 310

AUTHOR	Owen, Steven V.; Feldhusen, John F.
TITLE	Using Performance Data Gathered at Several Stages of Achievement in Predicting Subsequent Performance.
INSTITUTION	National Council on Measurement in Education, East Lansing, Mich.; Purdue Univ., Lafayette, Ind.
PUB DATE	Mar 70
NOTE	23p.; Paper presented at the annual meeting of the National Council on Measurement in Education, Minneapolis, Minnesota, March, 1970
EDRS PRICE	EDRS Price MF-$0.25 HC-$1.25
DESCRIPTORS	*Academic Achievement, Achievement, Evaluation, Measurement, Nursing, *Performance, Performance Criteria, Prediction, *Predictive Ability (Testing), *Predictive Measurement, Student Personnel Services, *Success Factors

ABSTRACT

This study compares the effectiveness of three models of multivariate prediction for academic success in identifying the criterion variance of achievement in nursing education. The first model involves the use of an optimum set of predictors and one equation derived from a regression analysis on first semester grade average in predicting the subsequent semester indices. In the second model, an optimum set of predictors and a new equation are derived for each semester average, individually. For the third model, each semester's optimum set and equation are determined individually, but prior semester averages are not included. In the first model predictive efficiency declines sharply from semester to semester. The second model predictions remained stable over the four semesters. Statistical comparison between the second and third models shows that the inclusion of prior semester averages in the prediction battery strongly increased the predictive efficiency beyond that level attained for each sample's first semester average. (Author/MC)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

ED040406

Using Performance Data Gathered at Several Stages of Achievement in Predicting Subsequent Performance

Steven V. Owen and John F. Feldhusen, Purdue University

The deficit in nursing services has never been as crucial as it presently appears, and numerous forecasts predict an ever-increasing shortage (Flint and Spensley, 1969). Cornelius (1968) estimated the national deficit of nurses to be about 125,000. Flint and Spensley (1969), reviewing several hundred nursing manpower studies, concluded that since the nursing shortage is so apparent, it is time to stop studying the size of the shortage and to start doing more about it. One possible remedy is to increase the number of nursing school graduates. This is a superficial suggestion, however, which fails to take account of the many problems inherent in trying to increase training services.

Ostlund (1965) reported that at least a third of entering nursing students eventually drop out of nursing school. It does not seem an economical or efficient solution, then, for nursing schools to simply increase enrollment, since to do so does not face the problem of attrition due to academic failure. If the attrition rate is so high, it will take at least three hundred new students to produce two hundred graduates. With the current competition among disciplines for talent this is not a promising approach.

Because attrition implies a waste of time, money, and human resources, nursing schools have focused attention on two procedures intended to reduce the dropout rate: better screening and selection of students, and, remediation

A paper presented at the annual meeting of the National Council on Measurement in Education, Minneapolis, Minnesota, March 1970.

CG005310

of academic difficulties once students are enrolled. The prediction of academic success is a tool implicit in both of these procedures. Prediction of success is the *sine qua non* in student selection, because the goal of student admission is to ensure that those who matriculate also graduate. For academic remediation, prediction has not served the roal for which is capable. Too often, remediation begins after low performance is a *fait accompli* for the student. There is also reason to suspect that, having experienced academic failure, a student's prospects of returning to an adequate performance level are complicated by the possibilities of lower self-image, lower expectancy of her success by both student and faculty, less peer approval, a lower level of motivation, and negative attitudes toward school (Alexander, 1968). If accurate prediction of the need for remediation can be make, and if remediation is promptly undertaken, then some of the supposed concomitants of failure might be prevented.

Objectives

This research was directed toward prediction of academic achievement following admission; it was intended to fit within a scheme of remediation. That is, once predictions have been formulated, students whose predicted grades are low may be given some systematic attention. The aim of the present research was to compare the effectiveness of three models of multivariate prediction of academic success in identifying the criterion variance of achievement in nursing education.

Accuracy of assessment and prediction within academic settings has received considerable attention over the past few decades. However, despite the promise of technological and methodological advances in measurement, statistical analyses, and computer utilization, prediction of academic sucess has not fulfilled its apparent potential. Bloom and Peters (1961) reported that a 1934 review of twenty-three prediction studies boasted a median simple correlation of .55 between high school grades and college grades. A 1957 review of fifteen studies, according to Bloom and Peters, showed a median simple correlation of .41 between the same two variables. While this apparent regression may not be representative of the state of prediction, there is a possibility that academic variables have become more complex and perhaps more difficult to relate to each other.

With the recent recognition of the utility of multivariate prediction techniques, accuracy of prediction has been somewhat increased. Primarily, this is because multivariate techniques permit the combination of variables in predicting a criterion, and also allow the efficient elimination of those variables which do not aid in prediction. Lavin (1965), reviewing the multiple correlation technique in academic prediction of college achievement, reported a mean multiple correlation of .65 between such traditional predictors as ability and previous grades, and criteria of academic performance. In another review of prediction of college achievement, Bloom and Peters (1961) found multiple R's ranging from .55 to .65 with the use of academic predictors. Thus, the average amount of variance accounted for by academic predictor variables appears to be in the range of only about 30 to 40 percent.

While some researchers have relegated the untapped criterion variance to "error variance," others are more optimistic about identifying its components. Lavin (1965), for instance, suggested that the usefulness of most cognitive predictors may have peaked, and that we may profitably embark on a search for non-academic predictors of school achievement, to build upon that "cognitive peak." Lavin proposed that such variables as creativity, motivation, biographical data, and personality traits be investigated for their contributions to the prediction of academic success. Taylor et al. (1963), reviewing prediction studies in nursing education, strongly supported Lavin's position on the possibilities of non-cognitive predictor variables.

This research incorporates certain non-cognitive variables in the forms of biographical data, measures of creativity, and anxiety.

Prediction Models

Several approaches have been used in the development of academic prediction systems. Three of these approaches will be reviewed and compared in this research. The three approaches to prediction will hereafter be referred to as Model I, Model II, and Model III.

Model I represents the typical counseling or admissions assessment of student background and ability, for the purpose of predicting whether the student is likely to achieve staisfactorily in subsequent semesters. The admissions assessment is ordinarily concerned with a commom measure of success such as grades or standardized test scores, and the predictions, while not statistical, are made with the same predictor variables, from semester to semester.

While there are few reports of the success of academic prediction with this model, the student attrition is so bad in most schools that it seems clear that there is a need to seek alternative approaches to prediction. Sawyer (1966), reviewing clinical and statistical prediction systems, showed clearly that clinical prediction is rarely, if ever, more accurate than statistical prediction. With this evidence, it seems likely that a statistical version of Model I should produce predictions at least as accurately as the traditional admissions predictions.

The advantage of Model I is that all predictions of subsequent student performance are done immediately following the recording of the first semester index. On the other hand, Plapp, Psathas, and Caputo (1965) present results of statistical predictions which convincingly show different sets of intellective variables predicting consecutive grade averages in nursing school. Accordingly, they attack the assumption that initial prediction batteries remain consistently precise throughout a series of successive student performance criteria.

Model I predictions of subsequent semesters are likely to decline in accuracy, because the variables which best predict the first semester index may not predict equally well those criteria further removed in time. This could also occur because the component abilities which constitute the fourth semester index may be different from those which make up the first semester average. For example, a reading comprehension score may identify a large portion of predictable variance of the first semester average, but be of little importance in predicting fourth semester average.

Model II. This model represents a statistical system of prediction in which multivariate techniques are employed in the prediction of a succession of achievement indices. That is, new analyses are run and new prediction equations developed for each of a series of criteria using a set of predictor variables, all of which are obtained prior to assessment of the first criterion.

Morman, Liddle, and Heywood (1965), for instance, used several personality scales to predict nursing student grades for two successive semesters. Kelleher, Kerr, and Melville (1968) used a battery of personality and achievement measures in predicting subsequent success of nursing aides. The criteria of success were defined as "survival in training program, final exam score, salary increment time, and length of job tenure." Owen, Feldhusen, and Thurston (1968) determined optimum sets of predictors for each of four semester averages in a nursing education program. While the optimum batteries varied for different semesters, the multiple correlations remained stable at about .65 over the semesters.

Model III. Model III builds upon Model II in that the optimum predictor sets and the prediction equation are determined individually for each semester. In addition, each successive performance index is predicted using that optimum set plus prior performance indices. For example, predictors for the third semester average would include the third semester optimum set plus first and second semester averages. The rationale for this model is simply that variables which are similar both in history and composition should predict one another better than variables further removed in time and different in structure. Bloom (1964) lends support to this notion with his argument that the best predictors of achievement will be obtained from performances in a similar environment. Since the college surroundings are a somewhat different educational environment than the student has been accustomed to in high school, it appears reasonable that prediction of college success from performance within college should be more accurate than prediction from earlier high school performance or from tests taken before he enters college. In other words, the use of performance data gathered at a later stage should improve prediction over a system which uses the same type of data gathered earlier. Also, the similarity of the additional predictors (previous semester averages) to the criterion should augument prediction. Obviously, this is because there is more common variance in similar variables than in dissimilar ones.

Burgess and Duffey (1969) found multiple correlations as high as .84 when they used the first two years of pre-nursing grades as part of a battery predicting the last two years' performance in a baccalureate program of nursing. They condluded that much tim., effort, and money could be saved by using previous GPA in predicting future grade averages.

This research was designed to answer the following question: Are there differences in predictive efficiency among three models for prediction of achievement in nursing education?

Procedure

The subjects were students who entered five associate degree schools of nursing in the years 1965 (first semester N=109), 1966 (first semester N=142), 1967 (first semester N=181), and 1968 (first semester N=197). These samples all suffered attrition in subsequent semesters. The associate degree school is a two-year program of nursing.

All predictor variables are shown in Table 1. The battery of predictor variables was composed of the following measures gathered on all subjects: age, two tests of anxiety (Sarason, 1961; Taylor, 1953); a verbal fluency measure (Christensen, Guilford, Merrifield, and Wilson, 1960); a creativity self-report measure, with several subscales; two tests of short term memory (French, Ekstrom, and Price, 1963); father's and mother's occupational and educational status; SAT verbal and quantitative scores; high school cumulative grade average; high school grades in English, mathematics and science; high school graduation percentile rank; and, prior years of education. The 1968 entrants had scores on all variables mentioned above, and, in addition, the Nelson-Denny Reading Test (three subscores) (Nelson and Denny, 1960), the Guilford-Zimmerman Temperament Survey (ten subscores) (Guilford and Zimmerman, 1949), marital status, previous experience, the grade average the student expected to receive.

The criteria were each of four semester grade averages for those students enrolled as freshmen in the years 1965, 1966, and 1967. For 1968 freshmen, only the first two semesters were used as criteria, since those students had not yet completed their final year.

Model I. For Model I, an optimum set of predictors was determined for the criterion first semester grade average, by means of a stepwise multiple regression procedure. All of the predictor variables were assessed prior to or at the time of admission to a nursing education program. The optimum set is defined as that combination of variables which predicts the criterion with a minimum standard error of estimate. Equivalently, each predictor variable of the optimum set has an F value above 1.00. ~~significant at the .01 level~~.

After the optimum set was found, that same set of variables (and the resultant regression equation) was used to predict all subsequent semester averages.

Model II. As in Model I, the same set of predictors was used and optimum sets of predictors were found for the first semester grade index. While Model I involved the use of the same optimum set for subsequent semester predictions, in Model II optimum sets were determined for each semester individually with a multiple regression analysis. Thus, the only relationships that optimum sets have with each other is that they were derived from the same battery of predictors.

Model III. Again, optimum sets of predictors were determined for each individual semester. However, previous semester grade averages were included in the optimum sets. In other words, the Model III optimum set of predictors for the second semester average was identical to the Model II optimum set, except that the model III set included the first semester average as a predictor variable.

All three models involved a stepwise multiple regression method of prediction commonly known as a "buildup" procedure. The ultimate outcome of the buildup technique is a multiple correlation between all significant predictors and a criterion.

The format of the buildup method is a stepwise process which, in the first stage, computes a simple correlation between the criterion and that variable which shows the strongest relationship with the criterion. At the next stage, a second variable is added to the first variable, and a multiple correlation is computed between the two predictors and the criterion. The determination of which variable is to be added to each stage is made by assessing the contribution of all remaining variables with the first variable partialed out. The second strongest predictor, then, is the one which has the highest part correlation with the criterion, after the first predictor has been partialed out.

This stepwise procedure continues until the contribution of the remaining variables fails to meet a prespecified F value. ~~level of significance~~. With this technique it is possible to derive not only the multiple correlation among predictors and criterion, but also in descending order, the relative contribution of each significant predictor variable in identifying criterion variance.

Cross-validations were done by first combining the 1965, 1966, and 1967 samples and using random halves of that total sample. Prediction equations were developed for the first half of the sample on each individual semester.

ERIC

These equations were used to predict achievement for the second half of the sample. Finally, those predicted averages were correlated with the subjects' actual grade averages. Model I results were not cross-validated because it is assumed in the model that the students' actual grades would not have been obtained at the time of the development of the optimum set of predictors.

It was necessary to cross-validate the 1968 sample separately from the combined sample, since a slightly different battery of predictor variables had been used with the 1968 entrants. The same cross-validation technique described above was employed with this separate sample.

Results

Results for the 1965 and 1966 samples, for all three models, are presented in Table 2. Table 3 shows results for the 1967 and 1968 entrants.

Model I. The optimum set of predictor variables was derived for the first semester grade index. That same optimum set was then used to predict the grade averages for the subsequent three semesters. (In truth, all three models start out at the same point, the first semester index. Thus, each model's first semester optimum predictors develop the same multiple correlation.) In Models II and III, of course, the predictions were developed independently for each of the four semesters.

Using the regression equation for the optimum set of predictors of first semester average, the subsequent three semester predictions showed significant, but steadily declining correlations. While the multiple correlation for predicting the first semester index averaged .67 for all four years, the average correlation for the second semester index dropped to .59. For the third semester predictions, the average correlation was .40, and for the fourth semester predictions, .36. The averaged correlations are presented in Table 4.

The predictions of 1966 entrants suffered the most severe decline, dropping from a first semester R of .62 to a fourth semester correlation of .26. Further, the correlations fluctuated from sample to sample. For instance, while the 1965 and 1967 samples' fourth semester correlations were respectively .40 and .42, the 1966 sample only produced a correlation of .26 for the same semester.

Model II. Optimum predictor batteries were generated for each individual semester. Results indicate that, while the average multiple correlations tended to decline over the four semesters, the drop was not nearly so abrupt as in Model I. When multiple regression was carried out on semester two, the average multiple R was .66, as compared to the average first semester R of .67. For third and fourth semesters, the average R's were .56 and .59, respectively.

The largest discrepancy in results between the samples was in the fourth semester, between the 1965 and 1966 entrants. The fourth semester R for the 1965 sample was .65 and for the 1966 sample, .56. Thus, the four semester predictions in Model II appeared to be fairly stable, both over the semesters, and accross the years. In addition, the correlations were generally higher than those found for Model I.

Results of Model II cross-validation are shown in Table 6. Since the prediction battery was altered for the 1968 sample, separate cross-validations were done for the 1968 entrants. Results indicate a minimal amount of shrinkage from the validation analyses, with a single exception. The shrinkage of the 1966, 1967, and 1968 entrants in the third semester predictions was substantial;

the validation multiple correlation of .47 dropped to an r of .26 on cross-validation.

Model III. Starting with the optimum batteries which were derived for each semester in Model II, all previous semester averages were incorporated as new predictor variables in those batteries. New predictions were then made for the second, third, and fourth semester averages. Following the average first semester multiple correlation of .67, the average second, third, and fourth semester R's were, in order, .79, .70, and .73. It is seen here that the multiple correlations were actually augmented in the semesters following the first semester.

Since Model III built directly upon Model II, a statistical comparison between the two models was possible. The tests of differences between Model II and Model III predictions are presented in Table 5. Results show that the increments made beyond the Model II correlations were highly significant. For instance, a Model II optimum battery generated an R of .66 for the 1965 entrants' second semester; the Model III battery (the same set of predictors, plus the first semester average) produced an R of .78 for the same criterion. The difference between the R of .66 and the R of .78 was significant beyond the .001 level. The cross-validations for this model showed little shrinkage.

In summary, the inclusion of prior semester averages in the prediction battery strongly increased the predictive efficiency beyond that level attained for each sample's first semester average. In addition, the multiple R's derived for each semester were significantly greater than those R's obtained in Model II. Finally, the cross-validations indicate that the Model III predictions are quite stable.

Discussion and Conclusions

This research was concerned with the question: Are there differences among the three models in the efficiency of prediction of nursing school achievement? Model I involved the use of an optimum set of predictors and one equation derived from a regression analysis on first semester grade average in predicting the subsequent three semester indices. In Model II, an optimum set of predictors and a new equation was derived for each semester average individually. For Model III, each semester's optimum set and equation was again determined individually, but prior semester averages were included in the battery of predictor variables.

In Model I, it was apparent that predictive efficiency declined sharply from semester to semester. While the first semester mean multiple R in Model I was .67, the predictions of the next three semesters dropped off sharply to .36. Except as a moderate estimation of a student's subsequent performance in general, Model I type of predictions appear to have limited potential for use in identifying students who will need remedial instruction activities.

Model II predictions, done independently for each semester, remained stable over the four semesters. Essentially, with the use of Model II predictions, it is possible to take advantage of the shifting criterion variance and to go back to the original battery of predictors to select out new predictor variables or to assign new weights to previously used predictors in generating new equations for each semester. The Model II cross-validations indicated that the regression equations were moderately stable for the first, second, and fourth semesters. However, the cross-validations of the combined sample of 1965, 1966, and 1967 entrants suffered large shrinkage in the prediction of the third semester average. The cross-validations of the 1968 sample also showed large shrinkage. This may be due to the fact that, when a random hold-out sample is used for cross-validation, the sampling distribution of differences between the original and the various hold-out samples will vary from very small to very large. In the case of the original and hold-out samples for the analyses of 1968 data apparently the difference was very large. If random fluctuation is large, and it occurs with an important predictor variable, then the reliability of the regression equation would be weakened. Random fluctuations of scores is particularly likely to happen when the sample pool is small, as in the 1968 sample. On the whole, however, the Model II predictions seem to represent a quite useful technique for predicting students' need for remediation.

For Model III predictions, it was seen that the inclusion of previous semester averages in the predictor battery strongly augmented predictive efficiency. In all cases, the multiple correlations rose for the second, third, and fourth semesters beyond that level attained for the first semester. In comparisons between Model II and Model III predictions, significant differences were shown for each yearly sample and each semester in favor of Model III. In addition, most cross-validations of Model III regression equations showed an insignificant amount of shrinkage. The exception again occurred in the 1968 sample second semester cross-validation, where the sample was fairly small. While Model II seems to represent an adequate means of predicting school achievement, Model III predictions indicate a powerful increase in predictive efficiency over Model II.

While the levels of predictive efficiency attained in Model III seem quite promising, since as much as 74 percent of the criterion variance (R=.86) was accounted for in one of the predictions, most of the predictions still leave room for improvement.

Throughout this program of research which was initiated in 1964, predictor variables have been selected first on the basis of their theoretical relevance to the criteria of achievement in nursing education. After empirical testing, a number of the variables have been retained for further use while others have been dropped. Variables shown in Table 1 (1968 entrants) constitute the current battery which has survived theoretical and empirical testing. Among those dropped because of poor predictive value or because they were too costly or impractical to secure are the following: the Nurse Attitudes Inventory (Thurston and Brunclik, 1965) and the Nursing Sentence Completions (Thurston and Brunclik, 1964), two standardized tests developed specifically for prediction of achievement in nursing education; and two tests of short term memory (Memory I: first names, Memory II: numbers for objects. French, Ekstrom, and Price, 1963).

Continuing efforts are being made to increase the efficiency of the original battery and to tap the "unpredictable" portion of criterion variance. Various non-cognitive and personality measures are now being tested for their usefulness in the prediction of achievement in nursing education. Further, cross-validated prediction batteries are currently being used in preliminary field-testing for aiding nursing student counselors and advisors in their guidance of students.

TABLE 1

Predictor Variables

1965, 1966, 1967 Entrants

1. Age (13)[1]
2. Taylor Anxiety Scale (6)
3. Sarason Test Anxiety Scale (9)
4. Creativity self-report scale; total scored responses (2)
5. Creativity self-report scale; factor 1 (9)
6. Creativity self-report scale; factor 2 (8)
7. Creativity self-report scale; factor 3 (5)
8. Creativity self-report scale; factor 4 (8)
9. Creativity self-report scale; total item analysis scored responses. (7)
10. Memory I: first names (1)
11. Memory II: numbers & objects (1)
12. Creativity self-report scale; total items correlated with a fluency measure. (6)
13. Creativity self-report scale; total items correlated with a flexibility measure. (3)
14. Occupation of father (5)
15. Education of father in years (3)
16. Occupation of mother (7)
17. Education of mother in years (6)
18. SAT-quantitative score (5)
19. SAT-verbal score (14)
20. High School Grades: English (8)
21. High School Grades: Math (11)
22. High School Grades: Science (7)
23. High School Grades: cumulative (11)
24. Prior education in years (7)
25. Percentile rank: high school graduating class. (5)

Model III Only

26. Semester one grade average
27. Semester two grade average
28. Semester three grade average

1968 Entrants

1. Age (2)
2. Taylor Anxiety scale (1)
3. Sarason Test Anxiety scale (1)
4. Guilford-Zimmerman: Factor G (0)
5. GZTS Factor R (3)
6. GZYS Factor A (3)
7. GZTS Factor S (1)
8. GZTS Factor E (0)
9. GZTS Factor O (2)
10. GZTS Factor F (1)
11. GZTS Factor T (1)
12. GZTS Factor P (0)
13. GZTS Factor M (0)
14. Nelson-Denny Reading Vocab. score (5)
15. Nelson-Denny Comprehension score (3)
16. Nelson-Denny Reading Rate score (4)
17. Guilford's Alternate Uses (4)
18. Occupation of father (5)
19. Education of father in years (0)
20. Occupation of mother (1)
21. Education of mother in years (0)
22. SAT-quantitative score (0)
23. SAT-verbal score (0)
24. High School Grades: English (0)
25. High School Grades: Math (1)
26. High School Grades: Science (0)
27. High School Grades: cumulative (3)
28. Prior education in years (1)
29. Percentile rank: high school graduating class (4)
30. Creativity self-report scale: factor 1 (2)
31. Creativity; factor 2 (3)
32. Creativity; factor 3 (2)
33. Creativity; total responses (0)
34. Creativity; item analysis (3)
35. Student's expected semester average (3)
36. Student's lowest acceptable semester average (5)
37. Student's marital status (5)
38. Previous related experience (yes,no) (0)

Model III Only

39. Semester one grade average
40. Semester two grade average
41. Semester three average

[1]For the 1965, 1966, and 1967 samples, 19 analyses were done; for the 1968 sample, 5 analyses were done. The number in parentheses represents the number of times each variable appeared in optimum sets of predictors.

TABLE 2

Correlation Coefficients for 1965 and 1966 Entrants

1965 Entrants

Model		Semester 1	Semester 2	Semester 3	Semester 4
I	N	109	94	77	68
	R	.68**	.52**[1]	.38**[1]	.40**[1]
II	N	109	94	77	68
	R	.68**	.66**	.63**	.65**
III	N	109	94	77	68
	R	.68**	.78**	.80**	.81**

1966 Entrants

Model		Semester 1	Semester 2	Semester 3	Semester 4
I	N	142	140	107	91
	R	.62**	.60**[1]	.33**[1]	.26*[1]
II	N	142	140	107	91
	R	.62**	.65**	.57**	.56**
III	N	142	140	107	91
	R	.62**	.86**	.70**	.75**

*$p < .05$.
**$p < .01$.

[1]These coefficients represent a simple correlation; the predicted semester averages, based on the first semester regression equation, were correlated with the actual semester averages.

TABLE 3

Correlation Coefficients for 1967 and 1968 Entrants

1967 Entrants

Model		Semester 1	Semester 2	Semester 3	Semester 4
I	N	181	165	122	104
	R	.67**	.59**[1]	.47**[1]	.42**[1]
II	N	181	165	122	104
	R	.67**	.62**	.49**	.56**
III	N	181	165	122	104
	R	.67**	.78**	.63**	.63**

1967 Entrants

Model		Semester 1	Semester 2	Semester 3	Semester 4
I	N	197	176		
	R	.69**	.63**[1]		
II	N	197	176	These two semesters had not been completed by the students at the time of writing.	
	R	.69**	.69**		
III	N	197	176		
	R	.69**	.74**		

**p < .01.

[1] These coefficients represent a simple product-moment correlation.

TABLE 4

Averaged Correlations for Four Semesters[1]

1965, 1966, 1967, and 1968 Entrants

	Model I	Model II	Model III
Semester 1	.67**	.67**	.67**
Semester 2	.59**[2]	.66**	.79**
Semester 3	.40*[2]	.56**	.70**
Semester 4	.36*[2]	.59**	.73**

* $p < .05$.
** $p < .01$.

[1]Averages were obtained by means of a weighted transformation to standard scores, from the formula given in McNemar (1962, p.140). The formula is

$$Z_{average} = \frac{(N_1 - 3)_{z1} + (N_2 - 3)_{z2} + (N_3 - 3)_{z3}}{(N_1 - 3) + (N_2 - 3) + (N_3 - 3)}$$

[2]Model I correlations, for second, third, and fourth semesters are simple correlations; Models II and III are multiple correlations.

ERIC

TABLE 5

Differences Between Model II and Model III*

Year	Semester	F	df	Significance
1965	2	45.76	1,80	.001
	3	22.88	2,66	.001
	4	17.32	2,51	.001
1966	2	167.15	1,127	.001
	3	15.54	2,96	.001
	4	14.60	3,77	.001
1967	2	88.66	1,155	.001
	3	14.69	2,113	.001
	4	4.28	3,93	.001
1968	2	24.97	1,158	.001

*Derived from a formula given in McNemar (1962, p.284). Degrees of freedom are $M_1 - M_2$ and $N - M_1 - 1$, where M = predictor variables and N = sample size.

TABLE 6

Cross-Validations of Model II and Model III

	Model II			
	1965, 1966, 1967 validation	Cross-validation	1968 validation	Cross-validation
Semester 1	R = .60 N = 227 13 variables	r = .51 N = 226	R = .79 N = 93 14 variables	r = .44 N = 84
Semester 2	R = .57 N = 200 5 variables	r = .53 N = 199	R = .77 N = 93 16 variables	r = .36 N = 89
Semester 3	R = .47 N = 154 10 variables	r = .26 N = 153		
Semester 4	R = .33 N = 137 5 variables	r = .27 N = 136		

	Model III			
	1965, 1966, 1967 validation	Cross-validation	1968 validation	Cross-validation
Semester 2	R = .75 N = 200 9 variables	r = .75 N = 194	R = .79 N = 93 14 variables	r = .66 N = 86
Semester 3	R = .59 N = 154 9 variables	r = .57 N = 149		
Semester 4	R = .63 N = 137 8 variables	r = .61 N = 129		

References

Alexander, E. D. The marking system and poor achievement. Teachers College Journal, 1964, 36: 110-113.

Bloom, B. S. Stability and Change in Human Characteristics. New York: Wiley, 1964.

Bloom, B. S. and Peters, F. R. The Use of Academic Prediction Scales. New York: Free Press of Glencoe, 1961.

Brandt, E. M., Hastie, B., and Schumann, D. Predicting success on state board exams. Nursing Research, 1966, 15 (1), 62-69.

Burgess, M. M. and Duffey, M. Prediction of success in a collegiate program of nursing. Nursing Research, 1969, 18(1), 51- .

Christensen, P. R., Guilford, J. P., Merrifield, P. R. and Wilson, R. C. Alternate Uses (form A). Beverly Hills, California: Sheridan Supply Company, 1960.

Cornelius, D. A. Changing winds. Minnesota Nursing Accent, 1968, 40: 3-4.

Flint, R. T. and Spensley, Karen C. Recent issues in nursing manpower; a review. Nursing Research, 1969, 18(3), 217-229.

French, J. W., Ekstrom, R. B., and Price, L. A. Manual for kit of reference tests for cognitive factors. Princeton, New Jersey: Educational Testing Service, 1963.

Guilford, J. P. and Zimmerman, W. S. The Guilford-Zimmerman Temperament Survey. Beverly Hills, California: Sheridan Supply Company, 1949.

Kelleher, E. J., Kerr, W. A., and Melville, N. T. The prediction of sub-professional nursing success. Personnel Psychology, 1968, 21(3), 379-388.

Lavin, D. E. The prediction of academic performance. New York: Russell Sage Foundation, 1965.

McNemar, Q. Psychological statistics. New York: John Wiley, 1962.

Michael, W. B., Haney, R., and Brown, S. W. The predictive validity of diversified measures relative to success in student nursing. Educational and Psychological Measurement, 1965, 25(2), 579-584.

Michael, W. B., Haney, R., and Jones, R. A. The predictive validity of selected aptitude and achievement measures and of three personality inventories in relation to nursing training criteria. Educational and Psychological Measurement, 1966, 26(4), 1035-1040.

Morman, R. R., Liddle, L. R., and Heywood, H. L. Prediction of academic achievement of nursing students. Nursing Research, 1965, 14(3), 227-230.

Nelson, M. J. and Denny, E. C. The Nelson-Denny Reading Test. New York: Houghton Mifflin Company, 1960.

Ostlund, L. A. An adjustment program for nursing students. Hospital Topics, 1965, 43, 58-61.

Owen, S. V., Feldhusen, J. F., and Thurston, J. R. Prediction of achievement in nursing education with cognitive, affective, and divergent thinking variables. A paper presented at the annual meeting of the National Council on Measurement and Evaluation. Los Angeles, 1969.

Plapp, J. M., Psathas, G., and Caputo, D. V. Intellective predictors of success in nursing school. Educational and Psychological Measurement, 1965, 25(2), 565-577.

Sarason, I. G. Test anxiety and intellectual performance of college students. Journal of Educational Psychology, 1961, 52, 201-206.

Sawyer, Jack. Measurement and prediction, clinical and statistical. Psychological Bulletin, 1966, 66(3), 178-200.

Taylor, C. W., et al. Selection and recruitment of nurses and nursing students. Salt Lake City: University of Utah Press, 1963.

Taylor, Janet L. A personality scale of manifest anxiety. Journal of Abnormal and Social Psychology, 1953, 48, 285-290.

Thurston, J. R. and Brunclik, Helen L. Nursing sentence completions. Eau Claire, Wisconsin: Nursing Research Associates, 1965.

Thurston, J. R. and Brunclik, Helen L. Nurse attitudes inventory. Eau Claire, Wisconsin: Nursing Research Associates, 1965.